2003

A Preliminary Linkage Map of the Hard Tick, *Ixodes Scapularis*

A. J. Ullmann
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado 80522, USA

J. Piesman
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado 80522, USA

M. C. Dolan
Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado 80522, USA

W. C. Black IV
Department of Microbiology, Colorado State University, Fort Collins, CO 80523-1677, USA

Insect Molecular Biology (2003) **12**(2), 201–210

A preliminary linkage map of the hard tick, *Ixodes scapularis*

A. J. Ullmann*, J. Piesman*, M. C. Dolan* and W. C. Black IV†

**Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado 80522, USA;*
†Department of Microbiology, Colorado State University, Fort Collins, CO 80523-1677, USA

Abstract

A linkage map of the *Ixodes scapularis* genome was constructed, based upon segregation amongst 127 loci. These included 84 random amplified polymorphic DNA (RAPD) markers, 32 Sequence-Tagged RAPD (STAR) markers, 5 cDNAs, and 5 microsatellites in 232 F_1 intercross progeny from a single, field-collected P_1 female. A preliminary linkage map of 616 cM was generated across 14 linkage groups with one marker every 10.8 cM. Assuming a genome size of ≈ 10^9 bp, the relationship of physical to genetic distance was found to be ≈ 300 kb/cM in the *I. scapularis* genome.

Keywords: *Ixodes scapularis*, linkage map, microsatellites, RAPD-SSCP, STARs, cDNA-SSCP.

Introduction

Ticks and the pathogens they transmit have long been an important cause of morbidity and mortality. Ticks are second only to mosquitoes as vectors of human and animal diseases (Sonenshine, 1991). *Ixodes scapularis* has become an important vector species because of its ability to transmit *Borrelia burgdorferi*, the causative agent of Lyme disease, in Eastern and midwestern North America (Spielman *et al.*, 1985). *I. scapularis* is also a vector of the causative agent of Human Granulocytic Ehrlichiosis (HGE) (Telford *et al.*, 1996) and *Babesia microti* the causative agent of human babesiosis (Spielman, 1976).

Much is known about the development, physiology, anatomy and vector competence of ticks, however, little is understood about the genetic basis of these traits. Karyotypes have been characterized for 103 of the ~830 known tick species (Oliver, 1977). However, a linkage map has never been produced for any of these or, for that matter, for any chelicerate arthropod. This is in part due to the minimal generation time required to establish a mapping family. For example, beginning with a fertilized P1 *Ixodes scapularis* mother, 1 week is required for her to feed to repletion. Oviposition requires another month, followed by 3 months for the larvae to hatch and for their cuticle to harden. An additional week is required for larval feeding. Nymphal development, hardening and feeding take an additional 3 months. Finally, adult development and feeding takes at least 3 more months. A further 2 weeks are required for mating to generate the F_1 offspring. This entire ≈ 43 week process is then repeated to generate F_2 progeny amongst which recombination frequencies are estimated. Thus at a minimum, over 1 year and 10 months is required to establish a mapping family.

Once an F_1 intercross family has been established, the modern paradigm for linkage mapping involves the simultaneous estimation of recombination amongst hundreds (Antolin *et al.*, 1996) or potentially thousands (Yasukochi, 1998) of highly polymorphic loci. This typically yields an intensive linkage map with one marker distributed every 5 cM. Many different genetic markers are currently in use for intensive linkage mapping. One class includes simple sequence repeats or microsatellites. Microsatellite based linkage maps have been generated for anopheline mosquitoes (Zheng *et al.*, 1996), mice (Dietrich *et al.*, 1992) and humans (Hudson *et al.*, 1995). Characterization of microsatellite markers in *I. scapularis* has already been described (Fagerberg *et al.*, 2001), but the abundance of and heterozygosity at microsatellite loci was generally low.

Random amplified polymorphic DNA amplified by the polymerase chain reaction (RAPD-PCR) provides another class of highly polymorphic markers. This technique uses a short (10 oligonucleotide) primer to amplify small regions of a genome (Williams *et al.*, 1990). The single primer is typically a unique but arbitrary sequence of nucleotides with a minimum G/C content of 60%. The low annealing temperatures used during PCR allow the primer to anneal to many, arbitrary regions of the genome. When the primer annealing sites in the genomic DNA template are inverted repeats

Received 12 September 2002; accepted after revision 23 December 2003.
Correspondence: Ms Amy J. Ullmann, Centers for Disease Control and Prevention, Division of Vector-Borne Infectious Diseases, PO Box 2087, Rampart Road, Foothills Campus, Fort Collins, CO 80522, USA. Tel.: +1 970 221 6435, fax: +1 970 221 6476, e-mail: aff1@cdc.gov.

(the 3′ ends of annealed primers face one another) and are ≤ 1.2 kb apart, then the intervening region is amplified. Alleles amplified by RAPD-PCR are typically dominant. Recessive alleles arise through mutations in or around the primer annealing site that cause the PCR process to fail. Genetic linkage maps utilizing RAPDs have been created for many insects including *Anopheles gambiae* (Dimopoulos *et al*., 1996), *Aedes aegypti* and a parasitic wasp, *Bracon hebetor* (Antolin *et al*., 1996), *Apis mellifera* (Hunt & Page, 1995), and *Bombyx mori* (Yasukochi, 1998). However, a great deal of the information needed to estimate recombination frequencies is lost when mapping loci with dominant alleles.

Single strand conformation polymorphism (SSCP) analysis of RAPD markers can reveal alternative, co-dominant alleles (Antolin *et al*., 1996). SSCP analysis is based on the principle that the electrophoretic mobilities of single-strand DNA molecules in non-denaturing gels are dependent upon both the size and shape of the fragments. Several stable structures or conformations can be formed when secondary base pairing occurs among nucleotides on a single DNA strand. The length, location and number of intrastrand base pairs determine the secondary and tertiary structure of a conformation. Intrastrand interactions are highly sensitive to the primary sequence of the molecule. The SSCP technique detects 99–100% of point mutations in DNA molecules 100–300 base pairs (bp) in length and at least 89% of mutations in molecules 300–450 bp in length (Orita *et al*., 1989; Hayashi, 1991; Hiss *et al*., 1994; Vidal-Puig & Moller, 1994).

Sequence-tagged RAPD loci (STARs) (Bosio *et al*., 2000) are another useful modification of RAPD markers. A RAPD-PCR band from a SSCP gel is cloned into a plasmid vector and sequenced. This sequence is then used to design targeted primers for PCR. STAR products are analysed on a second SSCP gel, and usually segregate as co-dominant markers.

The SSCP analysis of cDNA sequences provides another class of highly polymorphic markers for use in linkage mapping. Although these are unique, functional regions of a genome, polymorphisms are frequent and can be used to generate linkage maps. This method has been used to map cDNA makers in *Ae. aegypti* (Fulton *et al*., 2001).

In this paper we describe the use of microsatellites, RAPD-PCR, STAR and cDNA markers to generate a preliminary linkage map in an F_1 intercross family of *I. scapularis.* For the reasons described earlier, *c.* 2 years were required to generate the F_1 intercross mapping family. We describe our successes and failures in developing polymorphic markers and provide a preliminary 616 cM linkage map across 14 linkage groups, the 1 *N* number of chromosomes determined cytogenetically (Oliver, 1977). This provided a linkage map with one marker every 10.8 cM. However, counting markers that did not map to any of these 14 linkage groups, we estimate a much larger map size of 3165 cM, with an average resolution of only one marker every 55 cM.

Results

The P_1 generation began as a field-collected, fertilized female. The genotype of the P_1 male had therefore to be inferred at all loci. Recombination frequencies among all marker loci were analysed in 232 F_2 progeny.

Twenty RAPD primers (Table 1) amplified 84 polymorphic marker loci. There were no criteria for selecting RAPD primers, and they were chosen from an existing set in the laboratory, and primers that provided amplicons were utilized in the mapping analysis. Genotypes at 64 loci were in expected Mendelian ratios among the F_2 offspring. Sixty-three of these loci and D04.800a and b (bands ≈ 1 mm apart on the SSCP gel that did not completely co-segregate, Fig. 2) were extracted from SSCP gels, reamplified, cloned and sequenced to generate 65 primer pairs (Table 2). Of these, 52 STAR loci were polymorphic, and genotypes at 33 of these were in expected Mendelian ratios in the F_2 offspring. Alleles at 20 of the 33 STAR loci segregated as co-dominant markers. Examples of genotypes among the P_1 and F_1 parents and nine F_2 offspring appeared for three STAR loci in Fig. 1.

At D02.328ST the P_1 female had a unique genotype and the F_1 parents were heterozygous. All three genotypes were recovered in the F_2 offspring. Alleles at D12.425ST demonstrated a unique allele in the P_1 female that was seen in the F_1 male, while the other F_1 parent was heterozygous. Alleles at D18.353BST demonstrated a unique

Table 1. List of RAPD primers and sequences utilized for mapping

Primer	Sequence
A09	GGGTAACGCC
A20	GTTGCGATCC
B15	GGAGGGTGTT
B18	CCACAGCAGT
B20	GGACCCTTAC
C19	GTTGCCAGCC
C01	TTCGAGCCAG
C04	CCGCATCTAC
D02	GGACCCAACC
D03	GTCGCCGTCA
D04	TCTGGTGAGG
D07	TTGGCACGGG
D08	GTGTGCCCCA
D12	CACCGTATCC
D13	GGGGTGAGCA
D16	AGGGCGTAAG
D17	TTTCCCACGG
D18	GAGAGCCAAC
D19	CTGGGGACTT
D20	ACCCGGTCAC

Table 2. List of 95 primer pairs for PCR amplification of 65 STAR, 20 cDNA and 10 microsatellite loci in *I. scapularis*

Locus name	GenBank Accession number	T_a	Length (bp)	Primer sequences
STARs				
A09.306ST	BZ385505	49	306	GGGTAACGCCAGGGTTTTCC
				GGGTAACGCCCGATGTATAG
A20.310ST	BZ592381	51	310	GTTGCGATCCCTCAGAGCGA
				GTTGCGATCCAACGAAGTTT
A20.390ST	BZ592382	57	390	GTTGCGATCCCAGGATATAC
				GTTGCGATCCTASAGCACAT
A20.517ST	BZ592383	59	517	GTTGCGATCCCTGCGCCTAT
				GTTGCGATCCAGGCGATCAC
B15.874ST	BZ385506	59	874	GGAGGGTGTTGCACAGTACA
				GGAGGGTGTTGGGGTGTGTG
B18.358ST	BZ385507	62	358	CCACAGCAGTCAAACCTTCT
				CCACAGCAGTAGTGATACTC
B18.653ST	BZ385508	62	653	CCACAGCAGTCGACCATGCG
				CCACAGCAGTAGTGATACTC
B20.361ST	BZ385509	62	361	GGACCCTTACGAGCGAAAGA
				GGACCCTTACCCCAACCAAA
C13.357ST	BZ385510	62	357	AAGCCTCGTCGGGGTAGAGA
				AAGCCTCGTCCTGCTGTTCT
C13.579ST	BZ385511	60	579	AAGCCTCGTCGTTAGACGTT
				AAGCCTCGTCCACTTTCCTT
C16.362ST	BZ385512	62	362	CACACTCCAGTCAGCATCAG
				CACACTCCAGGACCAATAGT
C16.432ST	BZ385513	46	432	CACACTCCAGGCAAAAAAAC
				CACACTCCAGCATTAGGGAA
C16.458ST	BZ385514	59	458	CACACTCCAGGCAGATGCAT
				CACACTCCAGCATCAAGTGC
C16.680ST	BZ385515	48	680	CACACTCCACCAATCCCCTA
				CACACTCCACGATACAAACG
C01.169ST	BZ385516	59	169	TTCGAGCCAGGGCAGGACCT
				TTCGAGCCAGCATTGATAAG
C01.170ST	BZ385517	59	170	TTCGAGCCAGCATTGAGAAG
				TTCGAGCCAGGGCAGGACCT
C04.323ST	BZ385518	60	323	CCGCATCTACGTACGATTGA
				CCGCATCTACGCCCGGTAGT
C04.331ST	BZ385519	57	331	CCGCATCTACATATACAATA
				CCGCATCTACCCCATCGACG
C04.345ST	BZ385520	57	345	CCGCATCTACGTTAGCAGTC
				CCGCATCTACACACAGTTTC
D02.328ST	BZ385392	50	328	GGACCCAACCCTATACCGCT
				GGACCCAACCTACGGAGGAG
D02.330ST	BZ385393	48	330	GGACCCAACCCTATACCGCT
				GGACCCAACCTACGGAGGAG
D02.460ST	BZ385394	45	460	GGACCCAACCATGTTTTAAA
				GGACCCAACCCCAAGAAATC
D02.463ST	BZ385395	48	463	GGACCCAACCCCCAAGAATC
				GGACCCAACCATGTTTTAAC
D02.464ST	BZ385396	48	464	GGACCCAACCCCAAGAATCC
				GGACCCAACCATGTTTTAAC
D02.466ST	BZ385397	48	466	GGACCCAACCCCAAGAAATC
				GGACCCAACCATGTTTTAAC
D02.470ST	BZ385398	50	470	GGACCCAACCATGTTTTAAC
				GGACCCAACCCCCAAGAATC
D02.472ST	BZ385399	48	472	GGACCCAACCATGTTTTAAA
				GGACCCAACCCCAAGAATCC
D04.410ST	BZ385400	45	410	TCTGGTGAGGGGGAAGAAGG
				TCTGGTGAGGCAACTACTGA
D04.534ST	BZ385401	45	534	TCTGGTGAGGCGTTCAAGGC
				TCTGGTGAGGAGGAATGTTC
D04.800AST	BZ85433–BZ85434	45	800	TCTGGTGAGGTTCGTGTCAA
				TCTGGTGAGGCGGGGCTCCA
D04.800BST	BZ85435–BZ85436	45	800	TCTGGTGAGGCGGGACTTCA
				TCTGGTGAGGTTCGTGTCAA
D07.292ST	BZ385402	45	292	TTGGCACGGGCACACAAATA
				TTGGCACGGGGCTGCAACAA
D07.434ST	BZ385403	50	434	TTGGCACGGGCAATCATTGC
				TTGGCACGGGACGCGGAAGC
D07.457ST	BZ385404	50	457	TTGGCACGGGCGCGTGGCTT
				TTGGCACGGGGGACAAAAGA
D07.649ST	BZ385405	50	649	TTGGCACGGGCAGCAGCCGG
				TTGGCACGGGTGGAGGTCGT

Table 2. (*continued*)

Locus name	GenBank Accession number	T_a	Length (bp)	Primer sequences
D08.459ST	BZ385406	50	459	GTGTGCCCCAGGACCTGGGT
				GTGTGCCCCACCGCAAAGAG
D08.463ST	BZ385407	50	463	GTGTGCCCCACCGCAAAGAG
				GTGTGCCCCAGGACCTGGGT
D12.425ST	BZ385408	50	425	CACCGTATCCAGGAAGTGTT
				CACCGTATCCTCGCAGAGTT
D12.617ST	BZ385409	36	617	CACCGTATCCCCCTAGCGAA
				CACCGTATCCAACAAAGGAT
D12.821ST	BZ385410	43	821	CACCGTATCCCGTTGGCCTA
				CACCGTATCCCCATGATTTC
D12.942ST	BZ385411	50	942	CACCGTATCCATGTGTAATC
				CACCGTATCCCCAGACTCGG
D17.459ST	BZ385412	45	459	TTTCCCACGGGCACGTACTA
				TTTCCCACGGCGTCATGACT
D17.984ST	BZ385413	48	984	TTTCCCACGGCAGGATAACT
				TTTCCCACGGATTAATCAAA
D17.1200ST	BZ385437–BZ85438	43	1200	TTTCCCACGGAACGACTCAC
				TTTCCCACGGCCAAACAAAG
D13.363ST	BZ385414	50	363	GGGGTGACGACCGACTGGC
				GGGGTGACGAGACAACGAAA
D13.443ST	BZ385415	50	443	GGGGTGACGATGGGTTGTGG
				GGGGTGACGACCAACCCGGG
D16.330ST	BZ385416	50	330	AGGGCGTAAGCAGAACCGTT
				AGGGCGTAAGCATCAGGTAT
D16.470ST	BZ385417	50	470	AGGGCGTAAGGGCCGATAGC
				AGGGCGTAAGTGCGTGTGTG
D16.521ST	BZ385418	50	521	AGGGCGTAAGCCTAACAAAA
				AGGGCGTAAGGGTAGTGTGT
D17.684ST	BZ385419	50	684	TTTCCCACGGCACCCTTTGT
				TTTCCCACGGGCCATTCCGC
D18.265ST	BZ385420	41	265	GAGAGCCAACACGTCAAATA
				GAGAGCCAACCTCACTGACC
D18.266ST	BZ385421	43	266	GAGAGCCAACCTCACTGACCT
				GAGAGCCAACACGTCAAAT
D18.284ST	BZ385422	39	284	GAGAGCCAACCTACTTGGCA
				GAGAGCCAACATGGAGAGGA
D18.353ST	BZ385423	39	353	GAGAGCCAACGGGTAAATAG
				GAGAGCCAACAAGAGACTAC
D18.366ST	BZ385424	50	366	GAGAGCCAACCCACCTTCAT
				GAGAGCCAACGCGCGGGCGA
D18.367ST	BZ385425	50	367	GAGAGCCAACGCGCGGGCGA
				GAGAGCCAACCCACCTACAT
D18.369ST	BZ385426	50	369	GAGAGCCAACCCACCTTCAT
				GAGAGCCAACCCGCGGGCGA
D18.826ST	BZ385427	50	826	GAGAGCCAACAATCAACCTC
				GAGAGCCAACGCCAAAGAAA
D19.382ST	BZ385428	50	382	CTGGGGACTTCAAACAACCC
				CTGGGGACTTGCCACGTCAG
D19.833ST	BZ385429	50	833	CTGGGGACTTCACAGAGGCA
				CTGGGGACTTCCGTTTGGAT
D20.284ST	BZ385430	50	284	ACCCGGTCACCAATTTGTAG
				ACCCGGTCACAGATGGAGAA
D20.294ST	BZ385431	50	294	ACCCGGTCACACGAGCGCTT
				ACCCGGTCACGTCTTCACTG
D20.817ST	BZ385432	50	817	ACCCGGTCACCAGCAGATTT
				ACCCGGTCACATCATTGTGT
cDNA				
Is218 (APOP-IN (AF333765) DAD-1-like protein [*Schistosoma japonicum*] (3.00E-41)*)				
	CA763761	60	209	GCCAGCGTAAACGAAACC
				GATGAAGCCCGACAGGAA
Is155 (ATP-S (mitochondrial ATP synthase [*Drosophila melanogaster*] (1.00E-25))				
	CA763762	52	195	CTGAAAGGGTCCCAGAAGA
				GGAATGGCACCGTAAAGC
Is108 (EF1-A (AF378368 elongation factor 1-α [*Coccidioides immitis*] (3.00E-34))				
	CA763763	62	196	TGTGGGCGTCATCAAGTC
				AAACAGAGTAACCAAACGAAAG
Is119 (ELF-2B (CAC08449 eukaryote initiation factor 2 β [*Gallus gallus*] (5.00E-57))				
	CA763764	58	301	CCTTCGCCAACTTCCTCG
				AAGCCAGACTTGATGCTCAC

Table 2. (*continued*)

Locus name	GenBank Accession number	T_a	Length (bp)	Primer sequences
Is211 (GST-1 (AF366931 glutathione S-transferase [*Boophilus microplus*]) (1.00E-30))				
	CA763765	53	218	GAAATAACGGTTGAGGGC
				GTAGGGAAGGTTGGGAAA
Is75 (GST-2 (JX0095 glutathione transferase (EC 2.5.1.18) b – guinea (pig) (1.00E-30))				
	CA763766	53	164	TCCTCGTTGAGATCCAGT
				GTTCCAGCAGTAATAGTCG
ISAC (AF270496 *I. scapularis* anticomplement protein)				
		56	188	TTTTGGCGATTTCGTTTC
				AATTCCTTCCCAGGTTGC
Is152 (ND(CG7580 gene product [*D. melanogaster*]putative NADH dehydrogenase)] (6.00E-27))				
	CA763767	53	251	GCACTTTGGGAACCTGAT
				TGGGCTGCTTCTACTTGTC
Is136 (QM (XP 141003 similar to ribosomal protein L10 [*Rattus norvegicus*]) (2.00E-49))				
	CA763768	62	184	AGGCCAAGGTGGATGAGTT
				TTGTTGATGCGGATGACG
Is30 (RPS12 (AF470687 40S ribosomal protein S12 [*Branchiostoma belcheri*]) (1.00E-49))				
	CA763769	62	214	GCTCATCTGTGCGTCCTG
				TCTCCTTGCCGTAGTCCTT
Is140 (RPS13 (AF116857 40S ribosomal protein S13 [*Cricetulus griseus*]) (3.00E-48))				
	CA763770	60	177	CCCAGTCGGCTCTGCCATAC
				TCTTGTTGCCCGTCACCC
Is45 (RPS14B (AF402822 40S ribosomal protein S14 [*Ictalurus punctatus*]) (3.00E-52))				
	CA763771	53	183	TTTGGCGTAGCCCACATA
				GACCTGCTTGCATTTCTCG
Is10 (RPS18) (AAN52390 ribosomal protein S18 [*Branchiostoma belcheri*]) (1.00E-65))				
	CA763772	55	206	TCGCCCTGACAGCCATTA
				CTGAACTTGCCGTCCTTG
Is66 (RPL12 (AAH08230 ribosomal protein L12 [*Homo sapiens*]) (1.00E-49))				
	CA763773	61	310	CCCAAGTTTGATCCCACG
				ATGTCAGGTTGCCGCTGT
Is149 (RPL17A) (AF395586 ribosomal protein L17/23 [*Spodoptera frugiperda*]) (1.00E-57))				
	CA763774	44	293	GCATCTCCCTCGGTCTTC
				ATTTCGCCTTTGTTGTTGA
Is171 (RPL40Ribosomal protein L40 ubiquitin extension protein [*D. melanogaster*] (2.00E-41))				
	CA763775	68	154	GACTGGTGGGTGGTGTCATT
				GGTCGCAGGTTGTTGGAG
Is28 (RPL44) (AAM94276 ribosomal protein L44 [*Chlamys farreri*]) (3.00E-49))				
	CA763776	45	250	ATGATAAGAAGGGCAAGGAT
				CTGGTGGATTAGAACTGGAT
Is27 (PRP(AF400199 ribosomal protein L37 [*Spodoptera frugiperda*] (6.00E-30))				
	CA763777	58	152	GCAGCTACCACATCCAAA
				CTGAAACGCCTCCAGACC
SALP16 (AF061845 *I. scapularis* salivary gland 16 kDa protein (salp16))				
		63	371	CAGTGAGACGGGAGCATC
				GTTCGGCACCTTCCTTTA
Is79 (UBQ (CAA44453 ubiquitin-conjugating enzyme [*D. melanogaster*] (2.00E-43))				
	CA763778	47	196	ACAGCAATGGCAGCATCT
				TTCCTGGTCCACTCCCTA
Microsatellites				
ISAG4	AF331743	41	250	AAGAAAATAAAGCGAACAAG
				ATAAGCAATTCATACGAGATAGT
ISAG25	AF331742	57	171	AAATGTCCGAACAGCCTTAT
				GCCCTTGAGTCTACCCACTA
ISAC4	AF331736–AF331739	49	170	GGGTCCCAACGATTGCTAAACCAG
				AAGCGTATCCGATTTGCCCTTCAT
ISAC8	AF331740	42	170	TCTTCCCGCTGCTGTCTCGTATTC
				GAGCTACCCCTTTCATCGTCTTCG
ISAC20	AF331734	42	150	AGAAACACGGAAGGAGAAAGGAGA
				AACTGTGCCAGATGGGAAAGAAGA
ISAC22	AF331735	44	200	CAGCTGGGCCCCTCCTTTTAATCC
				TATTGTAAGGCCAGTCGCCGCTGC
ISCAG12	AF331755	44	170	GAAGAACACCGAGCGAACCAGAAC
				GTGAGCTGAGGGTTGCTGTTGATG
ISGATA3	AF331752	47	150	AGTCCCCTCAGAGCGATTTTCA
				GGCCGCCAGTTTGATGGATA
ISGATA4	AF331753	41	200	CAGACAATGTCATTCAATCGCA
				CGCACAATGCAAAACAAATCTA
ISCTGY17	AF331745–AF331747	41	150	TTCTTGTTTTATTGGTTGGGTG
				AATGCAGGGTAAGTTGAGATTG

*Designates E score for cDNAs of unknown function. Information following the IS# designations on cDNAs indicate the results of a BlastX search on GenBank.

Table 3. Comparison of mapped and unmapped markers

Mapped markers	Unmapped markers	
ISAG25	ISAG4	D13.363ST
ISAC4	ISAC20	D16.330ST
ISAC8	ISAC22	D16.470ST*
ISGATA4	IACAG12	D16.521ST
ISCTGY17	ISGATA3	D17.459ST
A20.390ST	A09.306ST	D17.984ST*
C01.170ST	A20.310ST*	D17.1200ST*
C04.345ST	A20.517ST	D18.353ST
D02.328ST	B15.874ST*	D18.366ST
D02.330ST	B18.358ST	D18.367ST
D02.460ST	B18.653ST	D18.369ST
D04.800AST	B20.361ST	D18.826ST*
D04.800BST	C13.357ST	D19.382ST
D07.457ST	C13.579ST	D19.833ST
D13.443ST	C16.362ST*	D20.284ST
D17.684ST	C16.432ST	D20.294ST
D18.265ST	C16.458ST*	D20.817ST
D18.266ST	C16.680ST*	Salp16
D18.284ST	C01.169ST*	RPS18
ISAC	C01.170ST*	RPL17A
RPS12	C04.323ST	ELF-2B
RPL12	C04.331ST	RPS14B
EF1-A	D02.463ST	QM
	D02.464ST	RPL44
	D02.466ST	UBQ
	D02.470ST	RPS13
	D02.472ST	RPL40
	D04.410ST	APOP-IN
	D04.534ST*	GST-1
	D07.292ST	GST-2
	D07.434ST	PRP
	D07.649ST*	ND
	D08.459ST	ATP-S
	D08.463ST	
	D12.425ST	
	D12.617ST	
	D12.821ST*	
	D12.942ST*	

*Designates STAR loci not conforming to expected Mendelian ratios.

allele in the P_1 female that was seen in the F_1 male, while the other F_1 parent was homozygous. Genotypes at D02.328ST and D12.425ST were fully informative for mapping, and D18.353BST was partially informative. Alleles at the 17 remaining loci segregated as dominant/recessive markers, however, segregation patterns were fully informative in only 15 of these.

Fifty-six sequences from an *I. scapularis* salivary gland cDNA library (sequences kindly provided by Jose Ribeiro) were analysed for primer design. The primers were designed to amplify 250–400 bp fragments (optimal size for SSCP) in 20 of the genes (Table 2), with high matches to existing genes. However, ultimately only five of these could be mapped because the genomic DNA amplified by the primers was much larger than the anticipated size based upon the cDNA sequence. Amplified products from these loci were extracted from the polyacrylamide gels, reamplified and cloned into a plasmid vector for subsequent sequencing. As expected, we found that the cDNA sequence was interrupted by large introns, some up to 800 base pairs in size. SSCP analysis can only reliably reveal polymorphisms in fragments < 500 bp. We did not design internal primers from the genomic sequences.

Primers for 10 microsatellite loci (Fagerberg *et al*., 2001) are also listed in Table 2. However, only six of these could be mapped, one of which was not in the expected Mendelian ratios, and consequently only five microsatellite markers have been mapped. The remaining four were monomorphic, further supporting the observations of low heterozygosity found at *I. scapularis* microsatellite loci (Fagerberg *et al*., 2001).

A linkage map was generated using JOINMAP (Fig. 2). The linkage map contains only those loci at which genotypes were in Mendelian ratios. Fourteen linkage groups were identified by JMGRP at an LOD of 3.0. Above this LOD, markers began splitting away either singly or into separate groups. A total of 57 markers mapped across 616 cM on the 14 linkage groups with one marker every 10.8 cM. A total map distance of 3166 cM was estimated by making the assumption that the remaining 51 markers that did not co-segregate on to any of the 14 linkage groups were ≥ 50 cM from mapped markers (51 markers × 50 cM/marker) + 616 cM = 3166 (cM).

Linkage groups on the map were temporarily assigned numbers according to their relative lengths. These lengths could change if more markers were added to the map. Ultimately numbers will be assigned to chromosomes via physical mapping when a physical map for *I. scapularis* exists. When this is accomplished, chromosome numbers will be assigned according to overall chromosome length. Currently, the map shows that linkage groups 1 is four times longer than any of the other 13 linkage groups. We suspect that linkage group 1 corresponds to the sex chromosome, because it is known from cytogenetic work that in *Ixodes* the sex chromosome is 3–4-fold longer than the autosomes (Oliver, 1977), and when sex was mapped it was most closely linked to A09.583 on linkage group 1 with an LOD of 2.667.

The size of this linkage map suggests a very high recombination rate for the *I. scapularis* genome. The physical size of another ixodid tick genome, *Amblyomma americanum*, was previously determined to be 1.08 pg or 1.04×10^9 base pairs, and repetitive elements were in a long-period interspersion pattern (Palmer *et al*., 1994). Assuming that this is an accurate estimate of the size of the *I. scapularis* genome, then the relationship of physical to genetic distance is 267~329 kb/cM. Work is currently underway to determine the physical size and organization of the *I. scapularis* genome.

Discussion

Our experience in generating this preliminary linkage map demonstrates both promising aspects to and pitfalls for future work in tick genetics and genomics. Having tried

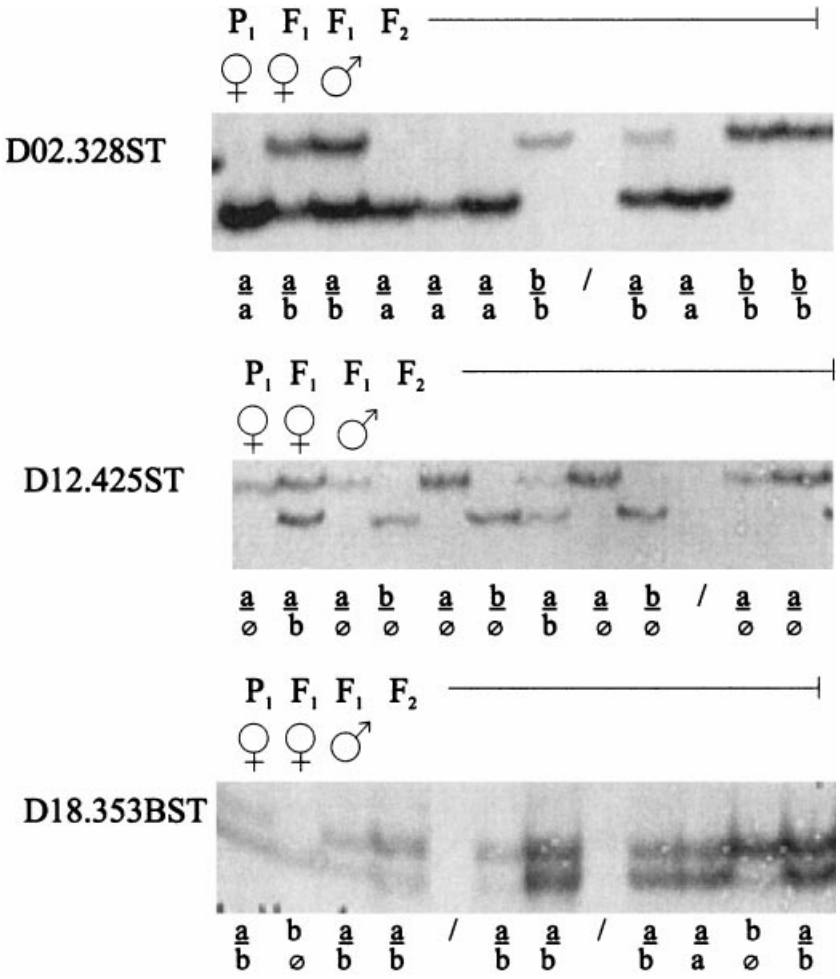

Figure 1. Genotypes at three STAR loci among the P_1 and F_1 parents and nine F_2 offspring. The Ø symbol designates null alleles.

conventional RAPDs, STARs, microsatellites and cDNAs, it appears that STARs are likely to be the marker of choice for expanding the *I. scapularis* linkage map and eventually integrating the linkage map into a physical map. Of the 65 STAR primer sets designed (Table 2), 52 were polymorphic and genotypes at 37 of these conformed to expected Mendelian ratios. The numbers of RAPD loci and the amount of heterozygosity at these loci were not as great as those detected in *Ae. aegypti* (Antolin *et al*., 1996). From the 20 RAPD primers, a total of 63 markers were mapped to the *I. scapularis* genome while 94 markers were generated with only 10 RAPD primers in an *Ae. aegypti* F_1 intercross family (Antolin *et al*., 1996). Reassociation kinetic studies of another ixodid tick, *Amblyomma americanum*, showed that the organization of repetitive DNA was of the long-period interspersion type (Palmer *et al*., 1994). We suspect that this is also the case for *I. scapularis.*

Fagerberg *et al*. (2001) had to use a specialized capture technique to identify a few microsatellite loci in *I. scapularis*. Furthermore, the heterozygosity at these few loci was low. We have no explanation for the low abundance and variability at microsatellite loci in *I. scapularis*. Sequences obtained from cDNA libraries were useful in generating PCR primers that amplified single nucleotide polymorphisms (SNPs) for linkage mapping in an *Ae. aegypti* F_1 intercross family (Fulton *et al*., 2001). In most cases the amplified regions were of the same size as the cDNA. However, this was not the case with cDNAs in the *I. scapularis* genome. In only 5 of the 20 cDNAs, for which primers were designed, were we able to recover a fragment that was sufficiently short for SSCP analysis. This appears to be due to the presence of large intervening introns. Taking the additional steps required to identify the locations and sequences of intron/exon boundaries will be very laborious and expensive.

Construction of the F_1 intercross family used in this mapping experiment was a time-consuming, laborious and expensive process. We used the majority of the genomic DNA from the parents and offspring to construct the map in Fig. 2. Normally, we would be unable to extend this map any further due to an inadequate amount of DNA, and would have to construct a new F_1 intercross family to map new markers.

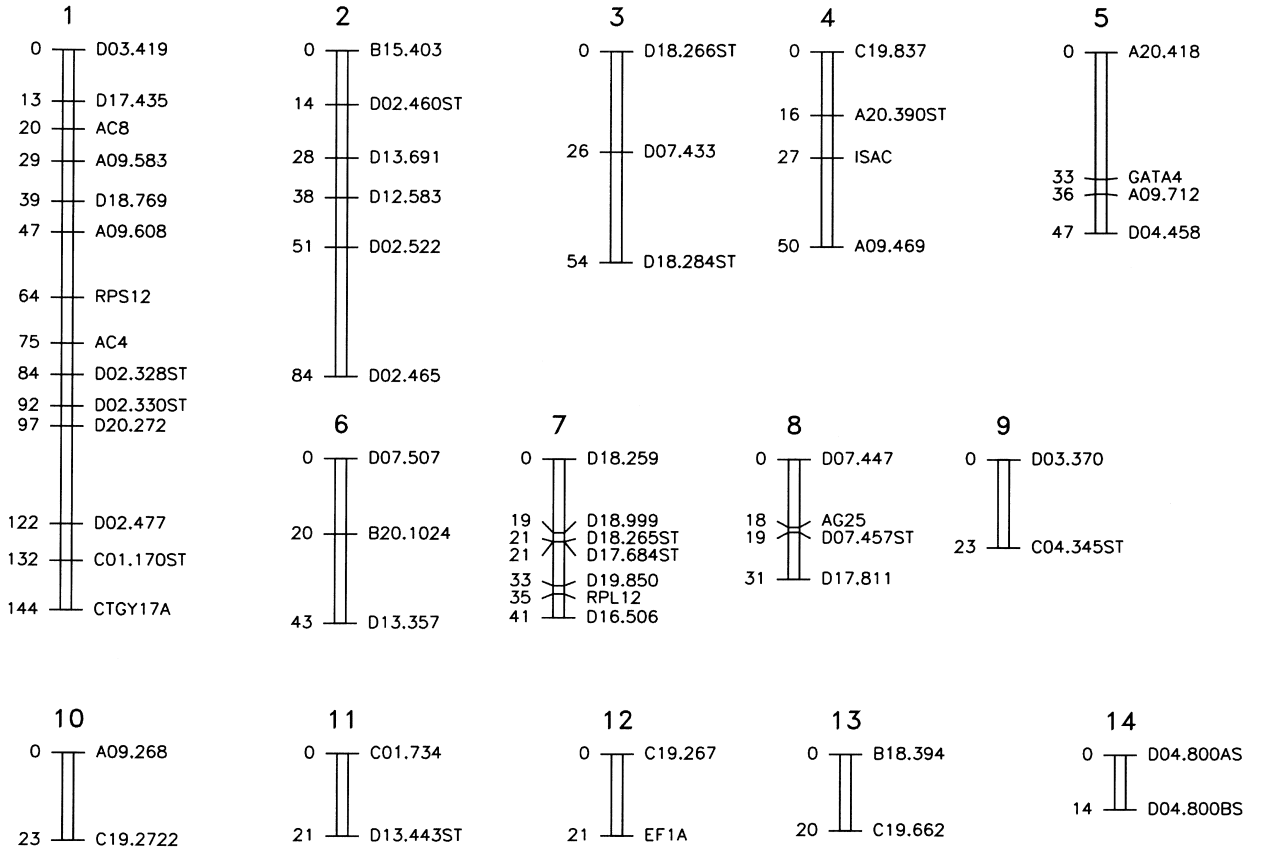

Figure 2. A preliminary linkage map of the *I. scapularis* genome including only those loci at which genotypes were in Mendelian ratios. Fourteen linkage groups were identified by JMGRP at an LOD of 3.0. A total of 57 markers map across 616 cM on 14 linkage groups with one marker every 10.8 cM. Linkage groups on the two maps were temporarily assigned numbers according to their relative linkage distances. These do not necessarily correspond to chromosomes numbered according to length cytogenetically (Oliver, 1977). RAPD markers are indicated by the name of the Operon primer, a period followed by the size of the amplified fragment. A STAR locus is indicated by a RAPD label followed by an 'ST'. cDNA loci follow the labels designated in Valenzuela *et al.*, 2002. Microsatellite loci are indicated by AC, AG, CTGY or GATA and the clone number.

However, we have recently amplified the whole genomes of the P_1, F_1 parents and F_2 offspring using multiple displacement amplification (MDA) (Gorrochotegui-Escalante & Black, 2003). The MDA-amplified genomic DNA can now be used to map additional markers as they become available.

We are in the process of performing the reassociation kinetics of the *I. scapularis* genome. However, assuming that the physical size of the *I. scapularis* genome is approximately the same size as the *A. americanum* genome (1.04×10^9 bp) (Palmer *et al.*, 1994) – perhaps an inappropriate assumption to make – our results suggest that there is a great deal of recombination in *I. scapularis*. A resolution of ≈ 300 kb/cM compares favourably to the average 1100 kb/cM resolution in *An. gambiae* (Dimopoulos *et al.*, 1996) or the 1000–3400 kb resolution for *Ae. aegypti* (Brown *et al.*, 2001). A map of this resolution has a higher probability of being successfully used in mapped-based positional cloning of candidate genes for vector competence to pathogens, host preference, insecticide resistance and other important characters in *I. scapularis*.

Experimental procedures

Mapping cross

All ticks were kept in Wheaton 8 ml sample vials (Wheaton, Millville, NJ) which were placed into glass desiccator jars with water in the bottom to maintain a high relative humidity. The desiccators were maintained in a Revco bioclimatic chamber (Revco Inc., Deerfield MI) at 21 °C, 95% humidity and received a photoperiod of 16 h: 8 h (light : dark). ICR outbred, pathogen-free mice were used for blood feedings (DVBID, CDC, Ft. Collins). Four fertilized P_1 females were collected from Bridgeport, Connecticut in autumn 1997, brought back to DVBID, and fed on rabbits. Eggs were laid by the P_1 females and hatched *c.* 4 weeks later to generate four initial families. After the eggs were collected from each P_1 mother, each female was frozen in a labelled cryotube at –20 °C in 70% ethanol. F_1 eggs from each female were hatched out and the larvae were blood fed on mice. One hundred replete larvae were place in Wheaton vials and *c.* 5 weeks after being fed, the larvae molted to nymphs. Once the cuticle hardened, the F_1 nymphs were blood fed. At this point the nymphs went into diapause for ≈ 8 months due to an unintentional drop in incubator temperature to 16 °C. When the incubator temperature was corrected, the ticks came out of diapause and development continued. When the fed

nymphs moulted to adults, eight single F_1 intercrosses were set up per family. Successful breeding pairs were established from P_1 female no. 2 (five pairs), and only a single pair from P_1 females nos. 1, 3 and 4. For each F_1 intercross, the females were isolated from the males, and single breeding pairs were made and subsequently fed individually in an ear bag on a rabbit ear.

Fed females were separated into labelled vials to collect F_2 eggs. These were laid from 4 to 6 weeks post-feeding. Each F_1 male was frozen in a labelled tube at −20 °C in 70% ethanol, and given the same number as his F_1 mate. Each F_1 female was frozen after collecting eggs. F_2 eggs were hatched and when the cuticle had hardened, blood fed on mice. After the larvae had moulted to nymphs, ≈ 4 weeks post-feed, and the cuticle had hardened, nymphs were fed on mice. The nymphs were reared to adults, and individually assigned a number that corresponded to the P_1 and F_1 female numbers. The F_2 adults were then frozen in a labelled tube at −20 °C in 70% ethanol.

Markers

Microsatellite loci were derived as described by Fagerberg *et al*. (2001). RAPD loci were amplified and run on SSCP gels as described in Black & DuTeau (1997). STAR loci were derived from polymorphic RAPD loci following exactly the procedure as described by Bosio *et al*. (2000). Primers for cDNA sequences from an *I. scapularis* salivary gland genomic library (kindly provided by Jose Ribeiro) were developed following the methodology of Fulton *et al*. (2001).

Map data analysis

Map distances were converted from recombination fractions to map units (cM) using the Kosambi mapping function (Kosambi, 1944). Offspring genotypes of F_1 and F_2 *I. scapularis* were entered into JoinMap 2.0 (Stam & van Oojin, 1995). Because the numbers of parental, recombinant and uninformative genotypes, and the segregation ratios among the F_2 offspring may differ for each locus, JoinMap allows each marker to be analysed according to parental genotypes. Initially, a threshold recombination fraction of 0.499 and a log odds density (LOD) score of 3.0 was used to group markers. The minimal LOD was then increased to 6.0 in increments of 0.1 to monitor the rate at which markers left individual linkage groups. Increasing the minimum LOD in increments tested how robust the linkage groups in the genotype data set are. DrawMap (van Ooijen, 1994) was used to plot a linkage map from the recombination frequencies generated by JoinMap.

References

Antolin, M.F., Bosio, C.F., Cotton, J., Sweeney, W., Strand, M.R. and Black, W.C.T. (1996) Intensive linkage mapping in a wasp (*Bracon hebetor*) and a mosquito (*Aedes aegypti*) with single-strand conformation polymorphism analysis of random amplified polymorphic DNA markers. *Genetics* **143**: 1727–1738.

Black, W.C.T. and DuTeau, N.M. (1997) *The Molecular Biology of Insect Disease Vectors: a Methods Manual*, pp. 361–373. Chapman & Hall, New York.

Bosio, C.F., Fulton, R.E., Salasek, M.L., Beaty, B.J. and Black, W.C.T. (2000) Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito *Aedes aegypti*. *Genetics* **156**: 687–698.

Brown, S.E., Severson, D.W., Smith, L.A. and Knudson, D.L. (2001) Integration of the *Aedes aegypti* mosquito genetic linkage and physical maps. *Genetics* **157**: 1299–1305.

Dietrich, W., Katz, H., Lincoln, S.E., Shin, H.S., Friedman, J., Dracopoli, N.C. and Lander, E.S. (1992) A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**: 423–447.

Dimopoulos, G., Zheng, L., Kumar, V., della Torre, A., Kafatos, F.C. and Louis, C. (1996) Integrated genetic map of *Anopheles gambiae*: use of RAPD polymorphisms for genetic, cytogenetic and STS landmarks. *Genetics* **143**: 953–960.

Fagerberg, A.J., Fulton, R.E. and Black, W.C. (2001) Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **10**: 225–236.

Fulton, R.E., Salasek, M.L., DuTeau, N.M. and Black, W.C.T. (2001) SSCP analysis of cDNA markers provides a dense linkage map of the *Aedes aegypti* genome. *Genetics* **158**: 715–726.

Gorrochotegui-Escalante, N. & Black, W.C. IV (2003) Amplifying whole insect genomes with multiple displacement amplification. *Insect Mol Biol* **12**: 195–200.

Hayashi, K. (1991) PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Meth Appl* **1**: 34–38.

Hiss, R.H., Norris, D.E., Dietrich, C.H., Whitcomb, R.F., West, D.F., Bosio, C.F., Kambhampati, S., Piesman, J., Antolin, M.F. and Black, W.C.T. (1994) Molecular taxonomy using single-strand conformation polymorphism (SSCP) analysis of mitochondrial ribosomal DNA genes. *Insect Mol Biol* **3**: 171–182.

Hudson, T.J., Stein, L.D., Gerety, S.S., Ma, J., Castle, A.B., Silva, J., Slonim, D.K., Baptista, R., Kruglyak, L., Xu, S.H., *et al.* (1995) An STS-based map of the human genome. *Science* **270**: 1945–1954.

Hunt, G.J. and Page, R.E. Jr (1995) Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. *Genetics* **139**: 1371–1382.

Kosambi, D.D. (1944) The estimation of map distances from recombination values. *Ann Eugen* **12**: 172–175.

Oliver, J.H. Jr (1977) Cytogenetics of mites and ticks. *Annu Rev Entomol* **22**: 407–429.

van Ooijen, J.W. (1994) DrawMap: a computer program for drawing genetic linkage maps. *J Hered* **85**: 66.

Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**: 874–879.

Palmer, M.J., Bantle, J.A., Guo, X. and Fargo, W.S. (1994) Genome size and organization in the ixodid tick *Amblyomma americanum* (L.). *Insect Mol Biol* **3**: 57–62.

Sonenshine, D.E. (1991) *Biology of Ticks*, Vol. 1. Oxford University Press, New York and Oxford.

Spielman, A. (1976) Human babesiosis on Nantucket Island: transmission by nymphal *Ixodes* ticks. *Am J Trop Med Hyg* **25**: 784–787.

Spielman, A., Wilson, M.L., Levine, J.F. and Piesman, J. (1985) Ecology of *Ixodes dammini*-borne human babesiosis and Lyme disease. *Annu Rev Entomol* **30**: 439–460.

Stam, P. and van Oojin, J.W. (1995) *JoinMap™ version 200: Software for the Calculation Linkage Maps*. CPRO-DLO, Wageningen, The Netherlands.

Telford, S.R., 3rd, Dawson, J.E., Katavolos, P., Warner, C.K., Kolbert, C.P. and Persing, D.H. (1996) Perpetuation of the agent of human granulocytic ehrlichiosis in a deer tick-rodent cycle. *Proc Natl Acad Sci U S A* **93**: 6209–6214.

Valenzuela, J.G., Francischetti, I.M., Pham, V.M., Garfield, M.K., Mather, T.N. and Ribeiro, J.M. (2002) Exploring the sialome of the tick *Ixodes scapularis*. *J Exp Biol* **205**: 2843–2864.

Vidal-Puig, A. and Moller, D.E. (1994) Comparative sensitivity of alternative single-strand conformation polymorphism (SSCP) methods. *Biotechniques* **17**: 490–492, 494, 496.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl Acids Res* **18**: 6531–6535.

Yasukochi, Y. (1998) A dense genetic map of the silkworm, *Bombyx mori*, covering all chromosomes based on 1018 molecular markers. *Genetics* **150**: 1513–1525.

Zheng, L., Benedict, M.Q., Cornel, A.J., Collins, F.H. and Kafatos, F.C. (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* **143**: 941–952.